7.6 Orthogonal Sets and Bases

Consider a set $S = \{u_1, u_2, \ldots, u_r\}$ of nonzero vectors in an inner product space V. S is called *orthogonal* if each pair of vectors in S are orthogonal, and S is called *orthonormal* if S is orthogonal and each vector in S has unit length. That is,

(i) **Orthogonal:** $\langle u_i, u_j \rangle = 0 \quad \text{for } i \neq j$

(ii) **Orthonormal:** $\langle u_i, u_j \rangle = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}$

Normalizing an orthogonal set S refers to the process of multiplying each vector in S by the reciprocal of its length in order to transform S into an orthonormal set of vectors.

The following theorems apply.

THEOREM 7.5: Suppose S is an orthogonal set of nonzero vectors. Then S is linearly independent.

THEOREM 7.6: (Pythagoras) Suppose $\{u_1, u_2, \ldots, u_r\}$ is an orthogonal set of vectors. Then

$$\|u_1 + u_2 + \cdots + u_r\|^2 = \|u_1\|^2 + \|u_2\|^2 + \cdots + \|u_r\|^2$$

These theorems are proved in Problems 7.15 and 7.16, respectively. Here we prove the Pythagorean theorem in the special and familiar case for two vectors. Specifically, suppose $\langle u, v \rangle = 0$. Then

$$\|u + v\|^2 = \langle u + v, \ u + v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle = \langle u, u \rangle + \langle v, v \rangle = \|u\|^2 + \|v\|^2$$

which gives our result.

EXAMPLE 7.9

(a) Let $E = \{e_1, e_2, e_3\} = \{(1, 0, 0), \ (0, 1, 0), \ (0, 0, 1)\}$ be the usual basis of Euclidean space $\mathbf{R}^3$. It is clear that

$$\langle e_1, e_2 \rangle = \langle e_1, e_3 \rangle = \langle e_2, e_3 \rangle = 0 \qquad \text{and} \qquad \langle e_1, e_1 \rangle = \langle e_2, e_2 \rangle = \langle e_3, e_3 \rangle = 1$$

Namely, E is an orthonormal basis of $\mathbf{R}^3$. More generally, the usual basis of $\mathbf{R}^n$ is orthonormal for every n.

(b) Let $V = C[-\pi, \pi]$ be the vector space of continuous functions on the interval $-\pi \leq t \leq \pi$ with inner product defined by $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)g(t)\, dt$. Then the following is a classical example of an orthogonal set in V:

$$\{1, \cos t, \cos 2t, \cos 3t, \ldots, \sin t, \sin 2t, \sin 3t, \ldots\}$$

This orthogonal set plays a fundamental role in the theory of Fourier series.

Orthogonal Basis and Linear Combinations, Fourier Coefficients

Let S consist of the following three vectors in $\mathbf{R}^3$:

$$u_1 = (1, 2, 1), \qquad u_2 = (2, 1, -4), \qquad u_3 = (3, -2, 1)$$

The reader can verify that the vectors are orthogonal; hence, they are linearly independent. Thus, S is an orthogonal basis of $\mathbf{R}^3$.

Suppose we want to write $v = (7, 1, 9)$ as a linear combination of u_1, u_2, u_3. First we set v as a linear combination of u_1, u_2, u_3 using unknowns x_1, x_2, x_3 as follows:

$$v = x_1 u_1 + x_2 u_2 + x_3 u_3 \qquad \text{or} \qquad (7, 1, 9) = x_1(1, 2, 1) + x_2(2, 1, -4) + x_3(3, -2, 1) \qquad (*)$$

We can proceed in two ways.

METHOD 1: Expand (*) (as in Chapter 3) to obtain the system

$$x_1 + 2x_2 + 3x_3 = 7, \qquad 2x_1 + x_2 - 2x_3 = 1, \qquad x_1 - 4x_2 + x3 = 7$$

Solve the system by Gaussian elimination to obtain $x_1 = 3$, $x_2 = -1$, $x_3 = 2$. Thus, $v = 3u_1 - u_2 + 2u_3$.

METHOD 2: (This method uses the fact that the basis vectors are orthogonal, and the arithmetic is much simpler.) If we take the inner product of each side of (*) with respect to u_i, we get

$$\langle v, u_i\rangle = \langle x_1u_2 + x_2u_2 + x_3u_3,\ \ u_i\rangle \qquad \text{or} \qquad \langle v, u_i\rangle = x_i\langle u_i, u_i\rangle \qquad \text{or} \qquad x_i = \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}$$

Here two terms drop out, because u_1, u_2, u_3 are orthogonal. Accordingly,

$$x_1 = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle} = \frac{7+2+9}{1+4+1} = \frac{18}{6} = 3, \qquad x_2 = \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle} = \frac{14+1-36}{4+1+16} = \frac{-21}{21} = -1$$

$$x_3 = \frac{\langle v, u_3\rangle}{\langle u_3, u_3\rangle} = \frac{21-2+9}{9+4+1} = \frac{28}{14} = 2$$

Thus, again, we get $v = 3u_1 - u_2 + 2u_3$.

The procedure in Method 2 is true in general. Namely, we have the following theorem (proved in Problem 7.17).

THEOREM 7.7: Let $\{u_1, u_2, \ldots, u_n\}$ be an orthogonal basis of V. Then, for any $v \in V$,

$$v = \frac{\langle v, u_1\rangle}{\langle u_1, u_1\rangle}u_1 + \frac{\langle v, u_2\rangle}{\langle u_2, u_2\rangle}u_2 + \cdots + \frac{\langle v, u_n\rangle}{\langle u_n, u_n\rangle}u_n$$

Remark: The scalar $k_i \equiv \dfrac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}$ is called the *Fourier coefficient* of v with respect to u_i, because it is analogous to a coefficient in the Fourier series of a function. This scalar also has a geometric interpretation, which is discussed below.

Projections

Let V be an inner product space. Suppose w is a given nonzero vector in V, and suppose v is another vector. We seek the ''projection of v along w,'' which, as indicated in Fig. 7-3(a), will be the multiple cw of w such that $v' = v - cw$ is orthogonal to w. This means

$$\langle v - cw,\ \ w\rangle = 0 \qquad \text{or} \qquad \langle v, w\rangle - c\langle w, w\rangle = 0 \qquad \text{or} \qquad c = \frac{\langle v, w\rangle}{\langle w, w\rangle}$$

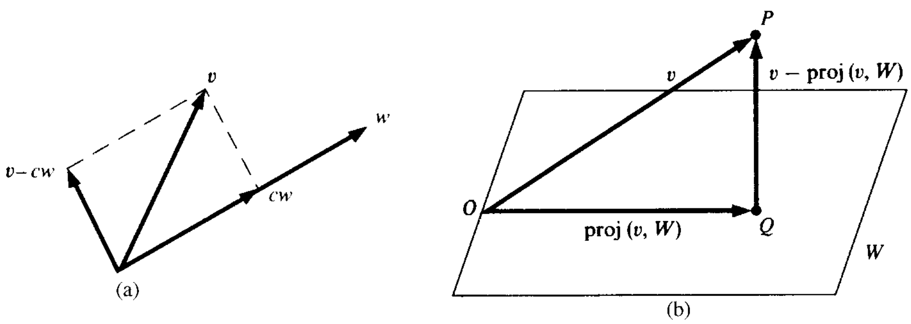

Figure 7-3

Accordingly, the *projection of v along w* is denoted and defined by

$$\text{proj}(v, w) = cw = \frac{\langle v, w\rangle}{\langle w, w\rangle}w$$

Such a scalar c is unique, and it is called the *Fourier coefficient* of v with respect to w or the *component* of v along w.

The above notion is generalized as follows (see Problem 7.25).

THEOREM 7.8: Suppose $w_1, w_2, \ldots, w_r$ form an orthogonal set of nonzero vectors in V. Let v be any vector in V. Define

$$v' = v - (c_1 w_1 + c_2 w_2 + \cdots + c_r w_r)$$

where

$$c_1 = \frac{\langle v, w_1 \rangle}{\langle w_1, w_1 \rangle}, \qquad c_2 = \frac{\langle v, w_2 \rangle}{\langle w_2, w_2 \rangle}, \qquad \ldots, \qquad c_r = \frac{\langle v, w_r \rangle}{\langle w_r, w_r \rangle}$$

Then v' is orthogonal to $w_1, w_2, \ldots, w_r$.

Note that each c_i in the above theorem is the component (Fourier coefficient) of v along the given w_i.

Remark: The notion of the projection of a vector $v \in V$ along a subspace W of V is defined as follows. By Theorem 7.4, $V = W \oplus W^{\perp}$. Hence, v may be expressed uniquely in the form

$$v = w + w', \qquad \text{where} \qquad w \in W \quad \text{and} \quad w' \in W^{\perp}$$

We define w to be the *projection of v along W*, and denote it by $\text{proj}(v, W)$, as pictured in Fig. 7-2(b). In particular, if $W = \text{span}(w_1, w_2, \ldots, w_r)$, where the w_i form an orthogonal set, then

$$\text{proj}(v, W) = c_1 w_1 + c_2 w_2 + \cdots + c_r w_r$$

Here c_i is the component of v along w_i, as above.

7.7 Gram–Schmidt Orthogonalization Process

Suppose $\{v_1, v_2, \ldots, v_n\}$ is a basis of an inner product space V. One can use this basis to construct an orthogonal basis $\{w_1, w_2, \ldots, w_n\}$ of V as follows. Set

$$w_1 = v_1$$

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1$$

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$w_n = v_n - \frac{\langle v_n, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_n, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 - \cdots - \frac{\langle v_n, w_{n-1} \rangle}{\langle w_{n-1}, w_{n-1} \rangle} w_{n-1}$$

In other words, for $k = 2, 3, \ldots, n$, we define

$$w_k = v_k - c_{k1} w_1 - c_{k2} w_2 - \cdots - c_{k,k-1} w_{k-1}$$

where $c_{ki} = \langle v_k, w_i \rangle / \langle w_i, w_i \rangle$ is the component of v_k along w_i. By Theorem 7.8, each w_k is orthogonal to the preceeding w's. Thus, $w_1, w_2, \ldots, w_n$ form an orthogonal basis for V as claimed. Normalizing each w_i will then yield an orthonormal basis for V.

The above construction is known as the *Gram–Schmidt orthogonalization process*. The following remarks are in order.

Remark 1: Each vector w_k is a linear combination of v_k and the preceding w's. Hence, one can easily show, by induction, that each w_k is a linear combination of $v_1, v_2, \ldots, v_n$.

Remark 2: Because taking multiples of vectors does not affect orthogonality, it may be simpler in hand calculations to clear fractions in any new w_k, by multiplying w_k by an appropriate scalar, before obtaining the next w_{k+1}.

Remark 3: Suppose $u_1, u_2, \dots, u_r$ are linearly independent, and so they form a basis for $U = \text{span}(u_i)$. Applying the Gram–Schmidt orthogonalization process to the u's yields an orthogonal basis for U.

The following theorems (proved in Problems 7.26 and 7.27) use the above algorithm and remarks.

THEOREM 7.9: Let $\{v_1, v_2, \dots, v_n\}$ be any basis of an inner product space V. Then there exists an orthonormal basis $\{u_1, u_2, \dots, u_n\}$ of V such that the change-of-basis matrix from $\{v_i\}$ to $\{u_i\}$ is triangular; that is, for $k = 1, \dots, n$,

$$u_k = a_{k1} v_1 + a_{k2} v_2 + \cdots + a_{kk} v_k$$

THEOREM 7.10: Suppose $S = \{w_1, w_2, \dots, w_r\}$ is an orthogonal basis for a subspace W of a vector space V. Then one may extend S to an orthogonal basis for V; that is, one may find vectors $w_{r+1}, \dots, w_n$ such that $\{w_1, w_2, \dots, w_n\}$ is an orthogonal basis for V.

EXAMPLE 7.10 Apply the Gram–Schmidt orthogonalization process to find an orthogonal basis and then an orthonormal basis for the subspace U of $\mathbf{R}^4$ spanned by

$$v_1 = (1, 1, 1, 1), \qquad v_2 = (1, 2, 4, 5), \qquad v_3 = (1, -3, -4, -2)$$

(1) First set $w_1 = v_1 = (1, 1, 1, 1)$.

(2) Compute

$$v_2 - \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = v_2 - \frac{12}{4} w_1 = (-2, -1, 1, 2)$$

Set $w_2 = (-2, -1, 1, 2)$.

(3) Compute

$$v_3 - \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 - \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = v_3 - \frac{(-8)}{4} w_1 - \frac{(-7)}{10} w_2 = \left(\tfrac{8}{5}, -\tfrac{17}{10}, -\tfrac{13}{10}, \tfrac{7}{5}\right)$$

Clear fractions to obtain $w_3 = (-6, -17, -13, 14)$.

Thus, w_1, w_2, w_3 form an orthogonal basis for U. Normalize these vectors to obtain an orthonormal basis $\{u_1, u_2, u_3\}$ of U. We have $\|w_1\|^2 = 4$, $\|w_2\|^2 = 10$, $\|w_3\|^2 = 910$, so

$$u_1 = \frac{1}{2}(1, 1, 1, 1), \qquad u_2 = \frac{1}{\sqrt{10}}(-2, -1, 1, 2), \qquad u_3 = \frac{1}{\sqrt{910}}(16, -17, -13, 14)$$

EXAMPLE 7.11 Let V be the vector space of polynomials $f(t)$ with inner product $\langle f, g \rangle = \int_{-1}^{1} f(t) g(t)\, dt$. Apply the Gram–Schmidt orthogonalization process to $\{1, t, t^2, t^3\}$ to find an orthogonal basis $\{f_0, f_1, f_2, f_3\}$ with integer coefficients for $\mathbf{P}_3(t)$.

Here we use the fact that, for $r + s = n$,

$$\langle t^r, t^s \rangle = \int_{-1}^{1} t^n\, dt = \left.\frac{t^{n+1}}{n+1}\right|_{-1}^{1} = \begin{cases} 2/(n+1) & \text{when } n \text{ is even} \\ 0 & \text{when } n \text{ is odd} \end{cases}$$

(1) First set $f_0 = 1$.

(2) Compute $t = \dfrac{\langle t, 1 \rangle}{\langle 1, 1 \rangle}(1) = t - 0 = t$. Set $f_1 = t$.

(3) Compute

$$t^2 - \frac{\langle t^2, 1 \rangle}{\langle 1, 1 \rangle}(1) - \frac{\langle t^2, t \rangle}{\langle t, t \rangle}(t) = t^2 - \frac{\frac{2}{3}}{2}(1) + 0(t) = t^2 - \tfrac{1}{3}$$

Multiply by 3 to obtain $f_2 = 3t^2 = 1$.

(4) Compute

$$t^3 - \frac{\langle t^3, 1\rangle}{\langle 1, 1\rangle}(1) - \frac{\langle t^3, t\rangle}{\langle t, t\rangle}(t) - \frac{\langle t^3,\ 3t^2 - 1\rangle}{\langle 3t^2 - 1,\ 3t^2 - 1\rangle}(3t^2 - 1)$$

$$= t^3 - 0(1) - \frac{\frac{2}{5}}{\frac{2}{3}}(t) - 0(3t^2 - 1) = t^3 - \tfrac{3}{5}t$$

Multiply by 5 to obtain $f_3 = 5t^3 - 3t$.

Thus, $\{1,\ t,\ 3t^2 - 1,\ 5t^3 - 3t\}$ is the required orthogonal basis.

Remark: Normalizing the polynomials in Example 7.11 so that $p(1) = 1$ yields the polynomials

$$1,\ t,\ \tfrac{1}{2}(3t^2 - 1),\ \tfrac{1}{2}(5t^3 - 3t)$$

These are the first four *Legendre polynomials*, which appear in the study of differential equations.

7.8 Orthogonal and Positive Definite Matrices

This section discusses two types of matrices that are closely related to real inner product spaces V. Here vectors in $\mathbf{R}^n$ will be represented by column vectors. Thus, $\langle u, v\rangle = u^T v$ denotes the inner product in Euclidean space $\mathbf{R}^n$.

Orthogonal Matrices

A real matrix P is *orthogonal* if P is nonsingular and $P^{-1} = P^T$, or, in other words, if $PP^T = P^T P = I$. First we recall (Theorem 2.6) an important characterization of such matrices.

THEOREM 7.11: Let P be a real matrix. Then the following are equivalent: (a) P is orthogonal; (b) the rows of P form an orthonormal set; (c) the columns of P form an orthonormal set.

(This theorem is true only using the usual inner product on $\mathbf{R}^n$. It is not true if $\mathbf{R}^n$ is given any other inner product.)

EXAMPLE 7.12

(a) Let $P = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 2/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \end{bmatrix}$. The rows of P are orthogonal to each other and are unit vectors. Thus P is an orthogonal matrix.

(b) Let P be a 2×2 orthogonal matrix. Then, for some real number θ, we have

$$P = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \qquad \text{or} \qquad P = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

The following two theorems (proved in Problems 7.37 and 7.38) show important relationships between orthogonal matrices and orthonormal bases of a real inner product space V.

THEOREM 7.12: Suppose $E = \{e_i\}$ and $E' = \{e_i'\}$ are orthonormal bases of V. Let P be the change-of-basis matrix from the basis E to the basis E'. Then P is orthogonal.

THEOREM 7.13: Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of an inner product space V. Let $P = [a_{ij}]$ be an orthogonal matrix. Then the following n vectors form an orthonormal basis for V:

$$e_i' = a_{1i}e_1 + a_{2i}e_2 + \cdots + a_{ni}e_n, \qquad i = 1, 2, \ldots, n$$